MIS Quarterly

SPECIAL ISSUE

GO TO YOUTUBE AND CALL ME IN THE MORNING: USE OF SOCIAL MEDIA FOR CHRONIC CONDITIONS[1]

Xiao Liu
Department of Information Systems, W. P. Carey School of Business, Arizona State University,
Tempe, AZ 85287 U.S.A. {xiao.liu.10@asu.edu}

Bin Zhang
Eller College of Management, University of Arizona,
Tucson, AZ 85721-0108 U.S.A. {binzhang@arizona.edu}

Anjana Susarla
Eli Broad College of Business, Michigan State University,
East Lansing, MI 48824 U.S.A. {asusarla@msu.edu}

Rema Padman
Heinz College of Information Systems and Public Policy, Carnegie Mellon University,
Pittsburgh, PA 15213-3890 U.S.A. {rpadman@cmu.edu}

Video sharing social media platforms, such as YouTube, offer an effective way to deliver medical information. Few studies have identified evidence-backed digital therapeutics with technology-enabled interventions to improve the ease with which patients can retrieve medical information to manage chronic conditions. We propose an interdisciplinary lens that synthesizes deep learning methods with themes emphasized in Information Systems and Healthcare Informatics research to examine user engagement with encoded medical information in YouTube videos. We first use a bidirectional long short-term memory method to identify medical terms in videos and then classify videos based on whether they encode a high or low degree of medical information. We then employ principal component analysis on aggregate video data to discover three dimensions of collective engagement with videos: nonengagement, selective attention-driven engagement, and sustained attention-driven engagement. Videos with low medical information result in nonengagement; at the same time, videos with a greater amount of encoded medical information struggle to maintain sustained attention-driven engagement. Our study provides healthcare practitioners and policymakers with a nuanced understanding of how users engage with medical information in video format. Our research also contributes to enhancing current public health practices by promoting normative guidelines for educational video content enabling management of chronic conditions.

Keywords: Visual social media, healthcare informatics, patient self-care, chronic diseases, deep learning, digital therapeutics, bidirectional long short-term memory (BLSTM)

[1]The accepting senior editors for this paper were Indranil Bardhan, Hsinchun Chen, and Elena Karahanna.

Introduction

Chronic diseases, such as cardiovascular conditions, cancer, and diabetes, are among the most common and costly of all health problems, many with high mortality and morbidity rates (Bauer et al. 2014). Over 117 million people in the United States have been diagnosed with one or more chronic disease, accounting for 84% of all healthcare spending in 2012 (Riegel et al. 2012). Chronic disease self-management and preventive health programs are a central plank of healthcare reform in the United States (Adams 2010). These programs promote informed lifestyle choices, risk factor modification, and active patient self-management of chronic diseases in order to improve patient outcomes and to reduce cost (Hernandez-Tejada et al. 2012; Simmons et al. 2014). The success of such programs relies heavily on accessible medical information and patient-centered, personalized communication practices (Ruppert et al. 2016).

Medical information refers to information created by a healthcare provider or consumer in any form or medium about physical, mental, or behavioral health or condition, the provision of healthcare to an individual, or payment thereof (Fair Credit Reporting Act of 1992). Traditionally, patients received such information from clinicians. Since the advent of the Internet and social media, the increased availability of medical information on the Internet, supplemented with that from clinicians, has enabled patients to seek medical information when and where they need it to manage their conditions (Eysenbach and Jadad 2001; Eysenbach et al. 1999; Sood et al. 2011; Steinberg et al. 2010). From a macroeconomic perspective, patients' ability to engage with medical information would improve not only the utilization of resources but also the health and well-being of the overall population (Parker et al. 2018).

It needs to be recognized that for individuals to realize the benefits of accessible medical information requires a high level of participation and engagement (Jordan et al. 2008). Only 12% of adults have proficient health literacy, according to the National Assessment of Adult Literacy (Kutner et al. 2006), which implies that 9 out of 10 adults lack the ability to understand basic medical information and engage in self-care and chronic disease management. It has been posited that purely text-based medical information results in low patient attention, comprehension, recall, and adherence, especially for patients with low literacy levels (Moorhead et al. 2013). Therefore, it is crucial to design patient education materials that increase patient attention and participation.

The rise of user-generated medical information disseminated through YouTube, which is the largest video sharing social media platform, could bridge this gap by providing information in a rich, visual format that may be easier to comprehend and adhere to (Backinger et al. 2011; Chatzopoulou et al. 2010). YouTube hosts more than 100 million videos providing information on the pathogenesis, diagnosis, treatments, and prevention of various medical conditions (Keelan et al. 2007; Madathil et al. 2015; Pandey et al. 2010; Ruppert et al. 2017, Sood et al. 2011; Steinberg et al. 2010). For patients with chronic conditions who need complex medical information, healthcare advice in a video format may make it more understandable and actionable, improving the efficiency of care.

YouTube offers the potential of empowering patients to be active managers of their conditions, enhancing patient well-being and enabling shared decision-making between patients and providers, ultimately improving the utilization of resources in the context of chronic care (Adams 2010; Parker et al. 2018). YouTube offers a diversity of content from professional healthcare organizations to user-generated medical information on a variety of chronic conditions (Madathil et al. 2015). For example, if we are searching for diabetes-related information on YouTube, we will find not only videos from healthcare organizations and hospitals but also information provided by healthcare advocates and patients, such as the various stages of diabetes, nutrition suggestions, and actionable advice such as how to operate an insulin pump (Fernandez-Llatas et al. 2017). Professionally produced videos by healthcare organizations are usually available only for a small subset of symptoms while user-generated videos cover more details of the day-to-day management of chronic diseases. Videos created by patients, caregivers, and healthcare practitioners, such as therapists and dietitians, complement videos from health organizations by providing guidelines on aspects such as weight management, exercise, pain management, administering medications, and support needed to actively self-manage chronic conditions. Providers, as well as patients, could use YouTube to share medical information such as health promotions (Backinger et al. 2011) and as a channel for patient education (Sood et al. 2011; Steinberg et al. 2010).

According to surveys by the Pew Research Center (Fox and Purcell 2010), most patients obtain information about chronic conditions through online health information searches. This raises two questions about the role of YouTube in better informing patients regarding chronic conditions. The first issue regarding information obtained through YouTube is the tremendous variation in the quality of content, especially the quality of medical information encoded in the videos (Pant et al. 2012). Currently, there is no validated method to assess whether the medical information available on YouTube conforms to the best available clinical knowledge or guidelines (Winkler et al. 2000). YouTube also lacks a centralized reputation system or a contextual ranking of queries. The second is the tremendous heterogeneity in content consumption or experience of content on YouTube. It has been estimated that almost 70% of YouTube videos were only viewed once, while the top 20% of videos managed to gain more than half of all views (Peck et al. 2008). Scholars have recognized the relatively ephemeral popularity of user-generated content and fluctuations in daily popularity of videos (Wu and Huberman 2007). This influences the process by which users search and obtain information. Viewers searching for popular medical symptoms could easily be led to a globally popular, but completely irrelevant, video. Since YouTube's own ranking system emphasizes popularity, misleading videos may acquire popularity due to the power laws in online attention (Keselman et al. 2008; Syed-Abdul et al. 2013).

The heterogeneity of informational needs from patients exacerbates the gaps in health literacy, eroding the value of YouTube as an avenue for patient education in chronic care (Mackert et al. 2016; Powers et al. 2017). For instance, a viewer searching for a generic keyword such as diabetes would encounter a highly popular video, but one that could contain misleading medical information. In contrast, a keyword search on a focused term such as insulin resistance is more likely to encounter reliable medical information produced by a reputable organization such as the Mayo Clinic (Fernandez-Llatas et al. 2017). In other words, the lack of knowledge about medical information may lead to misleading videos being retrieved by YouTube (Kallinikos and Tempini 2014). Low health literacy hinders patients' ability to search for medical information online, making patients with limited health literacy even more vulnerable. Healthcare providers and government agencies have expressed caution about the adverse consequences of medical misinformation on the Internet (Lau et al. 2012; Powers et al. 2017; Winker et al. 2000).

Our paper aims to address the gaps in understanding (1) what medical information about chronic conditions is available on YouTube and (2) how users collectively engage with the medical information discovered on YouTube. Regarding the first, a method to quantify the medical content in a video needs to be scalable to a vast repository of user-generated content (UGC) on YouTube. One of the earliest studies on this topic conducted a content analysis to evaluate the accuracy of YouTube videos on immunization (Keelan et al. 2007). In such studies, researchers could only analyze a small sample of videos for a narrow set of conditions given the amount of resources required to get experts to view and evaluate videos. The question then is how to develop an automated approach that can recognize medical terms. Regarding the second question, one of the gaps identified by extant research is how users search for and interact with UGC in the context of healthcare (Madathil et al. 2015). It has been posited that the process by which users seek and interact with medical information is a complex cognitive activity (Madathil et al. 2015). There is a paucity of methods and interventions to evaluate informational needs on video-sharing social media. While prior research on engagement with social media has primarily examined individual-level engagement, YouTube offers a unique context where we can leverage metadata to study aggregate engagement across the population of users. We build on theories of collective attention (Wu and Huberman 2007) to examine engagement with medical information encoded in videos as a measure of how users collectively interact with content obtained on YouTube.

In this work, we propose an interdisciplinary lens that synthesizes methods from Computer Science with user engagement themes emphasized in Information Systems (IS) and the medical information needs of patients in healthcare informatics (HIS) research. Medical information is often contained in the video description text in the form of medical terms and the underlying semantic association between them. We employ deep learning methods (LeCun et al. 2015) to develop a bidirectional long short-term memory (BLSTM) recurrent neural network to extract medical terminology from the user-generated content on video sharing social media sites. We collect metadata from 19,873 videos on diabetes and employ deep learning methods to extract the medical terms in these videos. Using annotated data, we build a logistic regression based classifier to characterize the encoded medical information in a video's metadata. To the best of our knowledge, this is the first study to extract complex medical information from real-world video data on social media platforms using deep learning methods. Using propensity score matching, we analyze collective engagement of YouTube videos as a proxy for user acceptance of medical information.

Our study makes the following contributions. First, we build on state-of-the-art deep learning methods to design and implement a scalable, automated, and generalizable approach that breaks through the inability of expert-based methods to assess medical information in video data. Second, we combine machine learning with econometric methods and annotation using domain experts to evaluate the impact of medical information on collective engagement on video sharing platforms. Our approach relies on building on and synthesizing economics, machine learning, and statistical methods to automate the challenging task of determining medical information in videos and categorizing them into high and low information videos based on collective engagement. Synthesizing multiple approaches allows us to create a generalizable framework to identify and extract rich and unique insights for analysis that can be applied to understand other health conditions and context. Our study provides best practice guidelines regarding how individuals and health organizations should engage health consumers with educational videos for chronic care. Our approach highlights new insights on collective engagement categories learned from aggregate video usage data that can be extended to understand the dynamics of collective engagement with user-generated content or video content.

Theory Development

Conceptualizing Encoded Medical Information in Video Sharing Social Media

Until recently, medical information was available to only a fraction of experts. The advent of social media and user-generated content has democratized access to medical information about chronic conditions both through patient-centric online communities as well as social media advocates that create and broadcast content related to various medical conditions. Medical information is available through digitized health databases such as MedlinePlus.[2] Medical information relates to clinical practice guidelines, scientific research about diseases, as well as the context of a specific patient or disease (Hersh 2002). Consumers now have access to a vast repository of medical information in social media. For instance, sites such as PatientsLikeMe[3] gather and share the codification of complex medical information through the participation of patients (Kallinikos and Tempini 2014). This enables healthcare consumers to bypass the filters from healthcare providers and allows them to actively seek out medical information on their own. In effect, the Web could evolve into an easily navigable global medical knowledge base, searchable across languages and continents (Eysenbach et al. 1999).

Medical terminology is considered a language used to precisely describe the human body including its components, processes, medical conditions, and any procedures operating upon it (Fernandez-Llatas et al. 2017). While prior literature has examined how healthcare professionals such as clinicians benefit from access to encoded medical information (Wyatt 2001), this could also apply to information seeking by patients (Heathfield and Louw 1999). Likewise, it can be surmised that the availability of encoded medical information through online sources empowers patients to seek out treatment options and reduces information asymmetry between healthcare consumers and providers (Eysenbach and Jadad 2001). Wyatt (2001) defines codification as a means of "identifying, capturing, indexing and making available explicit knowledge" (p. 1). Patients retrieving videos on chronic conditions through YouTube can, therefore, access codified medical information in a video format.

Medical information in online videos is conveyed through medical terminology, which constitutes healthcare-related words such as diseases, conditions, procedures, symptoms, and treatments (Fernandez-Llatas et al. 2017). Prior studies have relied on the judgment of domain experts such as health professionals to evaluate online medical information (Backinger et al. 2011; Dawson et al. 2011). Content rated by an expert (such as medical or allied professional) is the most common approach to assess the videos focused on health education. Medical information can be assessed along several critical dimensions, such as content understandability by end users (Ruppert et al. 2017), volume of medical information, complexity of medical information provided (Stellefson et al. 2014), and so on. We focus on the volume of chronic care related information encoded in a video. The Unified Medical Language System (UMLS) developed by the National Library of Medicine (NLM) maps an exhaustive lexicon of medical terms as concepts (Bodenreider 2004). We use UMLS as our framework to define the construct of encoded medical information and employ deep learning methods to extract encoded medical information. Deep learning allows computational models, composed of multiple processing layers based on neural networks, to learn representations of data with multiple levels of abstraction (LeCun et al. 2015). The first question we need to address is:

> ***Research Question 1***. *How do we conceptualize encoded medical information in a YouTube video?*

Collective Attention and Engagement with Videos

To understand how viewers interested in chronic care interact with videos retrieved through YouTube, we rely on theories of engagement and collective attention (Oh et al. 2010;Wu and Huberman 2007). Research on patient engagement has primarily focused on patients' ability to view their health data, especially in the case of chronic diseases (Simmons et al. 2014), and to better interact with clinicians and the healthcare system (Carman et al. 2013). By contrast, what can be gleaned from medical UGC and social media that cannot be assessed by patient engagement surveys from hospitals is user engagement with health information.

Users typically encounter videos related to chronic healthcare conditions through searches on YouTube (Eysenbach and Jadad 2001). YouTube videos made by healthcare organi-

[2]https://medlineplus.gov/

[3]https://www.patientslikeme.com/

zations as well as other patients help patients by providing information about symptoms and health experiences. These types of interactions also allow patients to benefit from other patients' experiences with different types of treatments and drug regimens, enabling them to comprehend the risks and challenges involved (Mao et al. 2013), enhancing patients' ability to understand and manage their condition (Eysenbach et al. 2004). Engaging with videos from both providers' and patients' perspectives empowers patients with higher self-esteem, greater confidence about their treatment, and acceptance of their illness (Rutten et al. 2016), Engagement is, therefore, a valuable proxy for understanding patient empowerment in actively managing chronic conditions. How users engage with videos is also a good proxy for acceptance and use of healthcare information found on YouTube. Instead of individual engagement, we employ collective engagement as a construct denoting attention by many individuals to health related videos on YouTube.

Prior studies have conceptualized engagement primarily at an individual level as "a desirable—even essential—human response to computer-mediated activities" (Laurel 1993, p. 112). Researchers have posited user engagement to be a multi-phase process wherein an individual interacts with an online interface (O'Brien and Toms 2008). Engagement can be a property of the users, the medium, or both (O'Brien and Toms 2008). Sidner et al. (2005) refer to engagement as a process by which individuals start, maintain, and end their perceived connection to computer-mediated activities. A common theme in this research is the involvement, either real or perceived, of the user in producing, consuming, or disseminating information (Napoli 2011; Oh et al. 2010).

Studies of collective attention have posited that attention is the cognitive process of concentrating on a discrete aspect of information that explains the human behavior of engagement and the consumption of information (Wu and Huberman 2007). Attention affects the propagation of information in social media, determining the effectiveness of advertising and viral marketing (Asur et al. 2011). Studies have shown that dynamics on YouTube demonstrate a strong positive dependence on attention (Asur et al 2011; Crane and Sornette 2008; Szabo and Huberman 2007) . Based on this literature, our next research question is about identifying the types of collective engagement on YouTube.

> ***Research Question 2***. *What types of collective engagement can we conceptualize on YouTube?*

Patient Information Seeking and Engagement in Video Sharing Social Media

Patients with chronic conditions need to have functional health literacy (Hernandez-Tejada et al. 2012), defined as the degree to which individuals have "the capacity to obtain, process and understand basic medical information and services needed to make appropriate health decisions" (Institute of Medicine 2004, p. 4). Health literacy encompasses not only the ability to read and interpret care instructions but also the ability to understand health-related risk (Mackert et al. 2016). Individuals with inadequate health literacy are less likely to describe their symptoms accurately; this, in turn, reduces their ability to communicate with healthcare professionals, thereby lowering health outcomes (Williams et al. 2002). In the context of video sharing social media, users search for and process information in a manner different from what is usually assumed in studies of information retrieval (Eysenbach and Kohler 2002). The considerable information overload and difficulty of finding relevant information exacerbates the problems in enabling patients to use video sharing social media to improve self-care options for chronic conditions. In reviewing the evidence from observational studies, Chou et al. (2013) posit that user-generated health content on social media such as YouTube and Twitter is largely inconsistent with that prescribed from clinical guidelines or scientific evidence. The videos that are easy for users to understand and those characterized by user engagement could be those that lack specifics needed to facilitate self-care. However, it could still benefit practitioners to co-opt user-generated content and consider the perspective of intended audiences as co-creators of content in order to design health messages with accurate and appropriate content.

It has been posited that public attention on social media platforms follows a phase transition type of dynamic characterized by power laws (Crane and Sornette 2008). YouTube's own recommendation systems and herding effects play a role in which videos get popular, resulting in attention being skewed toward a few popular videos. A majority of videos do not get views, a few become popular due to word of mouth effects endogenous to YouTube (Crane and Sornette 2008), while some acquire popularity from external factors; such popularity could get magnified by the self-referential dynamics of viewership wherein popular videos are viewed more. The fluctuating dynamics of attention on YouTube could be accentuated by the needs of viewers seeking healthcare information, leading to substantial heterogeneity in the nature of information seeking and consequently the type of attention.

Users encountering medical information in YouTube videos may experience or consume the content in different ways. Patients using YouTube to learn more about chronic conditions could be highly engaged encountering relevant videos on YouTube. Consequently, such engagement might translate to patients feeling empowered, as well as enabling them with greater confidence about their treatment and acceptance of their illness. At the same time, one of the challenges with medical terminology used in a video is that it could cognitively overburden a user. Studies of medical communication find that users encountering complicated medical terms are unable to comprehend healthcare instructions (Williams et al. 2002) and could even fail to recognize clinical terminology (Powers et al. 2010). Users encountering encoded medical information in a video could be bewildered by the complexity of terminology, which leads to lower engagement with the video. That is, some users motivated by the desire to acquire knowledge may perceive encoded medical information in a favorable manner; others motivated by social support may resist the cognitive demands from viewing such knowledgeable content.

While different types of attention facilitate engagement, it is also possible that users detach from or do not engage in the medical information obtained from the video at all. Prior literature has conceptualized information seeking as distinct from seeking hedonic experiences or personal enjoyment (O'Brien and Toms 2010). In the context of YouTube, both types of behavior are present. The preexisting health literacy of users may explain the divergence in how different users engage with medical information encoded in a video. On the one hand, users who exhibit greater health literacy may value the depth of medical information encoded in a video and dedicate sustained attention to the video. They may develop a stronger connection and interact with the video through commenting, linking, etc. On the other hand, those who lack health literacy cannot interpret the medical information encoded in a video. Their subsequent behavior, such as the affect and emotion toward the video and further interaction, would be to disengage with the video altogether. Understanding the processes of collective engagement enables us to better comprehend the connection between medical information in videos and the acceptance of such information through YouTube videos. It may also shed light into how YouTube content contributors should deliver medical information and what type of information users would like to engage with and consume. Thus, our research question about medical information and collective engagement is

> ***Research Question 3***. *What is the relationship between medical information encoded in a video and collective engagement with the video?*

Data and Methods

We present our research approach in Figure 1. We collect YouTube videos related to diabetes care with the YouTube Data API and 200 diabetes-related search keywords. We choose diabetes-related videos as our research test bed because diabetes is among the most prevalent chronic diseases in the United States (CDC 2017). Patient education plays a critical role in empowering diabetes management and improving patient outcome (Powers et al. 2017; Simmons et al. 2014). We extract medical terminology from the videos' metadata using a deep learning method and measure the different levels of collective engagement using the metadata. A machine-learning model is devised to classify the encoded medical information in videos using the inputs from deep learning and video-level features. We utilize propensity score matching to assess the causal effect of medical information in videos on collective engagement. Each component will be elaborated in the following subsections.

Data Collection

To assess the medical information and collective engagement level in YouTube videos, we develop a data collection process consisting of two steps: YouTube metadata retrieval and YouTube video retrieval. In YouTube metadata retrieval, we utilize YouTube data API to collect the video metadata. In YouTube video retrieval, we download videos directly from YouTube.

YouTube Metadata Retrieval

YouTube videos collected for this study must be relevant to chronic disease patient care and accurately represent what YouTube searches returned to patients searching for disease-related information. When retrieving the videos, we simulate patients' information seeking process for diabetes-related videos. A patient may search for medical information with terms related to the disease on YouTube. We select the search keywords from one of the largest online health communities, DailyStrength.[4] In DailyStrength's Expert Answers forum, patients ask questions about diabetes education and management, while certified medical experts provide answers. The search keywords used by patients on the Expert Answers forum represent typical informational needs of diabetes patients. Thus, we collect these terms from all the questions in the Expert Answers forum and define them as search keywords to retrieve YouTube videos. Figure 2 demonstrates the data collection process.

We fetch the top 100 videos for each search term and store the ranking of returned videos and their metadata in a database for further analysis. The information we collect from the YouTube Data API includes channel ID (account name), pub-

[4] http://www.dailystrength.org

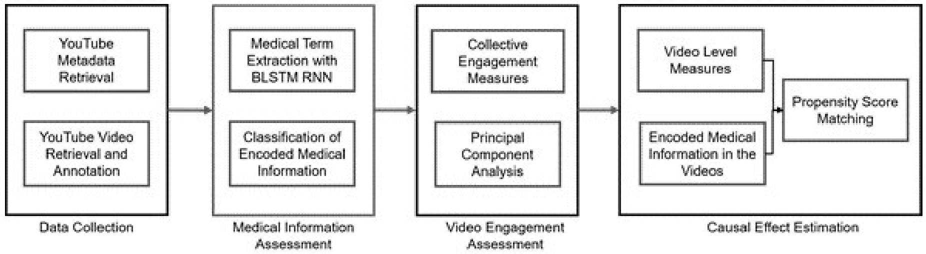

Figure 1. Proposed Research Approach

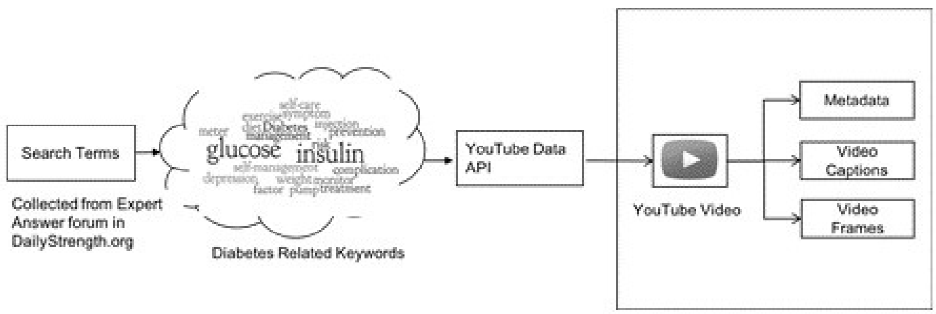

Figure 2. YouTube Video Data Collection Processes

lish time of the video, video title, video description, video tags, video duration, video definition, video caption availability, video rating, view count, like count, dislike count, and comment count. Video title, video description, and video tags are created by content owners when they upload the video. Furthermore, video usage data such as video rating, view count, like count, dislike count, and comment count reflects audience interaction and acceptance. In total, we collect 19,873 unique videos using over 200 search terms. These terms are available in Appendix A. The videos are contributed by both individual users and reputable healthcare organizations such as Mayo Clinic, American Diabetes Association, and American Nutrition Association.

YouTube Video Retrieval and Annotation

To perform supervised learning-based information extraction and classify encoded medical information in the video, we need an annotated video dataset. We select a subset of 600 videos that contain closed-captioning submitted by content contributors and download these videos and their captions. The videos and captions are used for video frame analysis and text analytics in robustness checks.

UMLS is used as a reference for medical terminology annotation. The open access and collaborative consumer health vocabulary (CHV)[5] is used to complement the existing framework of UMLS to connect informal, layperson words and phrases about health to medical terms used by health professionals. A total of 5,000 sentences are selected from the video description texts for annotation. Usage of medical terminologies in consumer vocabulary or standard medical format in these sentences are identified by referencing UMLS and CHV by two graduate research associates. The inter-rater reliability for annotating medical terms is 0.87. These two research associates independently viewed these 600 videos and ultimately labeled the videos for high or low medical information. The high or low medical information categorization is based on the annotators' experience as a consumer/viewer of the videos. Examples of high or low medical information videos are illustrated in Appendix C. A domain expert (a medical doctor) viewed all videos and consolidated the annotation results. The inter-rater reliability for classifying encoded medical information in videos is 0.92.

[5]https://www.nlm.nih.gov/research/umls/sourcereleasedocs/current/CHV/

Assessing Encoded Medical Information in YouTube Videos

The exponential growth in the volume of online videos on YouTube necessitates an automated approach to assessing the medical information in these videos at scale. We utilize a deep learning-based approach to recognize the medical terms in videos. Combining the extracted vocabulary of medical terms with other video features, we develop an automated method to classify the encoded medical information in videos.

Medical Term Extraction with Bidirectional LSTM Recurrent Neural Network Model

The characteristics of user-generated content such as informal words, and emoticons pose significant challenges to extracting medical terms. Most existing natural language processing tools that are trained on formal text (e.g., news articles) do not adapt well to informal, nonstandard language in UGC where medical terms could be expressed using consumer health vocabularies, slang, abbreviations, and even misspellings. Lexicon-based or knowledge-based methods are insufficient to detect the breadth of variations of medical terms used in social media due to the constantly evolving nature of online colloquial language. Conventional machine-learning methods such as conditional random fields (CRFs) treat words as discrete atomic symbols and predict a label for each word. The symbolic representation of words requires accurate input for training and prediction. Unfortunately, there are a large number of word variations in social media, and hence CRFs are not as successful in this text mining genre.

We utilize word embeddings, a vector representation, in our medical information extraction approach. Word embedding represents each word with its surrounding words in a vector (Mikolov et al. 2015). According to the distributional hypothesis, words with similar meanings occur with similar neighbors (Rubenstein et al. 1965). For instance, a 300-dimensional word embedding is composed of 300 words that are most likely to appear around the word of interest. This representation enables us to represent a medical term within its semantic context instead of isolating the symbolic term by itself and allows a machine learning model to learn the semantics of medical terminology more efficiently (relative to other methods) and improve the extraction performance.

In this study, we leverage a deep learning-based method, specifically, the bidirectional long short-term memory recurrent neural network, to extract medical terms from the video description and assess the level of medical information. Recurrent neural networks (RNNs) are a family of neural networks that operate on sequential data such as text and speech. In the context of medical entity extraction, RNNs take as input text a sequence of vectors (x_1, x_2, …, x_n) and return another output sequence (h_1, h_2, …, h_n). The output sequence represents entity labels of the sequence at every step in the input (Dos Santos and Zadrozny 2014). RNN models with word embedding input have achieved superior performance in many applications of natural language processing and understanding, such as parts-of-speech tagging (Dos Santos and Zadrozny 2014), named entity recognition (Baldwin et al. 2015), and machine translation (Sutskever et al. 2014). Although RNNs can learn long sequences (long input sentences) in theory, they fail to do so in practice and tend to be biased toward their most recent inputs in the sequence (Bengio et al. 1994). Long short-term memory networks (LSTMs) tackle this issue by incorporating a memory cell and have been shown to capture long-range dependencies. They do so with several gates that control the proportion of the input to give to the memory cell, and the proportion from the previous state to forget (Hochreiter and Schmidhuber 1997). LSTM RNNs have demonstrated superior performance for named entity recognition on noisy user-generated text (Baldwin et al. 2015; Lample et al. 2016). The algorithm presented in Table 1 describes the medical term extraction process in our study.

To identify medical terms, we train an embedding model using the Skip-gram method in Word2vec[6] on the entire collection of video descriptions. The implemented model obtains an array of semantic vectors, also known as the word embedding, containing predicted neighbor words of each word in the corpus. Word embedding enables us to represent consumer health vocabulary in social media textual content when its semantic information is similar to that of standard medical terms. Each unique word in the corpus was assigned a number. The Skip-gram method predicts what words are likely to co-occur with the word of interest. Below we detail the model. T is the number of unique words in the training corpus; c is the window size of the surrounding words. Words that appear within a distance of c are considered the surrounding words. w_{t+j} is the j^{th} word surrounding (before or after) w_t. Given a word w_t, the training objective is to maximize the average log probability:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\leq j\leq c, j\neq 0}\log p(w_t+j|w_t)$$

[6]Word2vec, designed by Google, is a set of methods that is widely used to produce word embeddings (https://www.tensorflow.org/tutorials/representation/word2vec).

Table 1. Medical Term Extraction

Algorithm 1. Medical Term Extraction from YouTube Videos

Input: A collection of sentences C.
A sentence $X \in C$ is represented as $[x_1, x_2, \ldots, x_n]$. x_i refers to the i^{th} word in the sentence X.

Output: A collection of labeled sentences.
Each sentence $X \in C$ is labeled as $L = [l_1, l_2, \ldots, l_n]$. l_i refers to the label for the i^{th} word in the sentence. $l_i \in \{MT, NA\}$. MT refers to a medical term. NA refers to a nonmedical term.

Procedures:
1. Train a word embedding model W with all $X \in C$.
2. For each $w_t \in V$, V is the vocabulary (unique words) in collection C, obtain the word embedding W_t of w_t.
3. Create an annotated data set $C' \subset C$. C' contains a collection of sentences S and their labels L'. Each w in C' is represented by its embedding W.
4. Build a BLSTM RNN model $L = R(S)$ with a collection of annotated dataset C'.
5. For each $X \in \{C\ \ C'\}$ (unlabeled sentences), apply the model R to obtain the labels for the sentence with $L' = R(X)$.

To avoid rare words negatively affecting the model performance, we prune the vocabulary by replacing the less frequent words with a unified symbol "UNK" (short for the unknown token). We keep the top 5,000 frequent words in their original form and replace the remaining words with UNK. Identifying medical terminologies in video descriptions can be considered a named entity recognition task. We devise a BLSTM model to extract medical terms from the user-generated video descriptions at the sentence level. The video descriptions are first tokenized into individual sentences and words. Figure 3 illustrates the BLSTM architecture for medical information extraction.

A BLSTM consists of two layers of LSTM units (i.e., LSTM1 and LSTM2 in Figure 3) that run in parallel: one follows the input sentence sequence, and the other follows the reverse of the input sentence sequence. BLSTM is a unique RNN architecture. The basic processing unit in the model is called the LSMT unit. Each LSTM unit contains an input gate $i^{(t)}$, a forget gate $f^{(t)}$, an output gate $o^{(t)}$, a memory cell $c^{(t)}$, and a hidden state $h^{(t)}$. $X^{(t)}$ is the word embedding input at time step t. Operator $\circ$ denotes element-wise multiplication. W, U, and b are the weight vectors of the gate parameters. The LSTM unit computes the output at each gate by iterating the following equations:

$$i^{(t)} = sigm(W_i x^{(t)} + U_i h^{(t-1)} + b_i)$$
$$f^{(t)} = sigm(W_f x^{(t)} + U_f h^{(t-1)} + b_f)$$
$$o^{(t)} = sigm(W_o x^{(t)} + U_o h^{(t-1)} + b_o)$$

$$u^{(t)} = \tanh(W_u x^{(t)} + U_u h^{(t-1)} + b_u)$$
$$c^{(t)} = i^{(t)} \circ u^{(t)} + f^{(t)} \circ c^{(t-1)}$$
$$h^{(t)} = o^{(t)} \circ \tanh(c^{(t)})$$

The forget gate controls how much information from each previous LSTM unit is forgotten, the input gate controls how much each unit is updated, and the output gate controls the exposure of the internal memory state. At each time step, the hidden state of the BLSTM is the concatenation of the forward and backward hidden states. This setup allows the hidden state to capture both the past and the future information. Then the word sequence is represented as an embedding sequence, which is passed to the BLSTM layer. Given an input sentence $(x_1, \ldots \ldots, x_n)$, our model computes the output label $(l_1, \ldots \ldots, l_n)$, regarding whether a given word in the sentence is a medical term, by iterating the following equations:

$$h^{(t)} = sigm(Wx^{(t)} + Uh^{(t-1)})$$
$$o^{(t)} = Vh^{(t)}$$

$sigm(\bullet)$ is the sigmoid function. U, W, and V are the weight vectors. At each time stage, our model takes the last hidden state $h^{(t-1)}$ and the current input $x^{(t)}$ to compute the current hidden state $h^{(t)}$, and it uses the current hidden state $h^{(t)}$ to compute the current output $o^{(t)}$. The current hidden state $h^{(t)}$ is further passed to the next iteration to calculate the next hidden state $h^{(t+1)}$. To reduce computational complexity, we train a 50-dimensional word embedding model, meaning

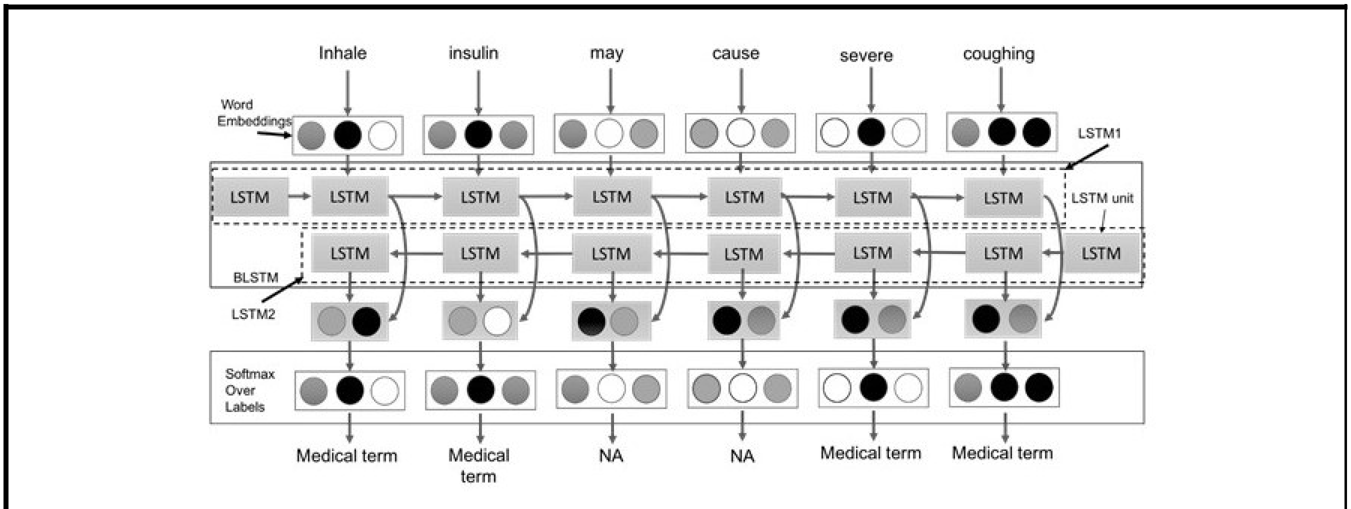

Figure 3. BLSTM Architecture for Medical Information Extraction

each word is converted to a 50-dimensional semantic vector. Instead of using a large hidden layer, we use 150 neurons in the BLSTM layer to avoid over-fitting. The outputs of the BLSTM layers are then processed to a Softmax classifier, which predicts the semantic type of each word in the input sentence.

Benchmarking and Performance Evaluation

We evaluate three popular approaches in information extraction as baseline methods to compare our proposed medical information extraction approach. The first baseline model is a lexicon-based approach using the UMLS as the lexicon resource. The UMLS offers a Java API—MetaMap—hat automatically tags medical terms in the National Library of Medicine's knowledge bases. Prior studies have utilized this tool to identify biomedical terms in texts (Bodenreider 2004). The second baseline model, CRF, is an undirected statistical graphical model that is widely considered a high-performance machine learning method for information extraction. We utilize CRFsuite[7] to train a CRF model to extract medical terms given new sentences. The third baseline model is the recurrent neural network model with word embedding representation. We adapt the implementation of RNN model from TensorFlow and train it on our corpus. The 600 annotated videos are divided into two parts: 80% for training and the remaining 20% for testing. Table 2 reports the performance evaluation of medical information extraction using precision (P), recall (R), and F-measure (F).

[7]http://www.chokkan.org/software/crfsuite

The results demonstrate that the BLSTM approach for medical information extraction outperforms the three baseline methods. The improvement is due to the vector representation of words within their context made possible through word embedding representation of medical terminologies expressed in consumer health vocabulary. The deep RNN further enhances sequence learning performance in social media texts. We then apply the BLSTM RNN model to the remaining unlabeled dataset. The BLSTM model extracted 218,603 medical terms, of which 6,514 are unique.

Other Video Level Information

In addition to the number of unique medical terms in the videos, other factors may influence users' perceived medical information level in the video. We combine heuristics-driven measures with automatically extracted medical terminology to assess the medical information of videos on YouTube. Table 3 shows the descriptive statistics of video level measures available from metadata. The descriptions of these measures are available in Table 4.

Assessing Collective Engagement with Videos

Collective Engagement Measures

The vast majority of engagement research on YouTube has relied on simple measures of raw exposure, with user behavior captured by gross metrics such as ratings, page views, and the like (Chatzopoulou et al. 2010; Siersdorfer et al. 2010).

Table 2. Performance of Medical Information Extraction

		Precision	*Recall*	*F-measure*
Baseline Model 1	UMLS	42.30%	21.90%	28.86%
Baseline Model 2	CRF	90.30%	74.50%	81.64%
Baseline Model 3	RNN	91.00%	88.50%	89.73%
BLSTM Approach	BLSTM RNN	**94.00%**	**91.80%**	**92.89%**

Table 3. Descriptive Statistics of the Video Level Measures

Numeric Variables	*Min*	Q_1	*Median*	*Mean*	Q_3	*Max*
# of unique words in description	0	19	48	80.9	119	444
# of words in description	0	22	64	147.5	195.8	1,005
Video duration (s)	1	181	353	677.7	712	9,716
# of unique medical terms in description	0	1	4	8.7	12	80
# of channel views	0	27,171	463,944	157.8	5,149,760	15,639,262,119
# of channel subscribers	0	74	1303	169,027	21,990	29,566,530
# of channel comments	0	0	0	157.8	2	104,922
# of channel video count	1	35	198	6,334	1000	1,013,672
# of channel average video view count	0	330	2,341	50,602	16,861	26,041,387
# of published days	1	274	721	1002	1,536	4,289
Categorical Variables	*Categories*					
Has title	True: 19,873			False: 0		
Has tags	True: 11,054			False: 8,819		
Has caption	True: 2,260			False: 17,613		
Content creator credibility	True: 3,034			False: 16,839		
Video definition	HD: 12,321			SD: 7,552		

Table 4. Descriptions of Video Level Measures

Video Level Measures	*Description*
# of words in the video description	Total number of words in the video description
# of unique words in the video description	Total number of unique words in the video description
Video duration	The total length of the video in the second
# of unique medical terms in video description	Total number of unique medical terms in video description
# of channel views	Total number of views the content contributor has
# of channel subscribers	Total number of subscribers the content contributor has
# of channel comments	Total number of comments the content contributor has
# of channel Video Count	Total number of videos the content contributor has
# of channel average video view count	Average video view count for the content contributor
# of published days	The number of days from content creation to the date of data collection
Has title	Whether the video has a title
Has tags	Whether the video has tags
Has caption	Whether the content contributor submits a caption together with the video
Content creator credibility	Whether a reputable healthcare organization manages the channel
Video definition	Video resolution (HD or SD)

Others have relied on eye tracking, physiological sensors, and follow-on task performance tests to measure these characteristics (Attfield et al. 2011). We propose a set of measures for collective engagement, building on prior research on user engagement. Prior studies of individual user engagement that have discussed characteristics associated with user engagement (Attfield et al. 2011), such as attention, affect, aesthetics, novelty, motivation, interest, feedback, control, and more (O'Brien and Toms 2008). Among the characteristics in prior literature, feedback, affect, motivation, and interest can be measured with publicly available YouTube metadata. These measures match with the prior theoretical framework of user engagement (O'Brien and Toms 2010). To measure users' feedback, we make use of the number of likes, number of dislikes, and number of comments. Within the psychological literature, affect has been thought of as the umbrella term that subsumes emotions, feelings, and sentiments (Fleckenstein 1991). Affect can be defined as positive and negative evaluations of an object, behavior, or idea with intensity and activity dimensions (Munezero et al. 2014). To measure user's affect, we examine the sentiment in comments with the Valence Aware Dictionary and Sentiment Reasoner (VADER; Hutto et al. 2014), an effective sentiment analysis tool for social media text. We use VADER to automatically classify the sentiment polarity of a given comment and compute the total number of positive, negative, and neutral comments for each video. Additionally, we develop two relevance measures to assess the motivation—first, the cosine similarity between search keywords and video title, and second, the cosine similarity between search keywords and video description. To evaluate the users' interest, we use the cosine similarity between user comments and the description of the video. Prior research has used the document relevance score to evaluate the quality of comments (Diakopoulos et al. 2015), while the video description relevance score can be considered as a reflection of user interest. We do find that the relevance score is correlated with comment quality. The cosine similarity function is defined in Appendix B. Table 5 shows the results of the collective engagement evaluation. The heterogeneity of user-generated content and the dynamic characteristics of YouTube platforms lead to the diverse values of these measures. The YouTube view counts range from 0 to over 1 billion. The number of likes is similarly skewed, ranging from 0 to over 1 million.

Table 6 examines the correlations among the collective engagement measures. View counts correlate with the number of dislikes, number of likes, and number of comments. The number of dislikes correlates with the number of likes and number of comments. The number of neutral comments correlates with the similarity score between comments and title, number of positive comments, and number of negative comments. The similarity score between comments and description is related to the similarity between comment and video title, number of positive comments, and number of negative comments. The similarity score is a reflection of users' interest in the video content. High interest is associated with high commenting activity.

Principal Component Analysis of Collective Engagement

As we observed the high correlations among collective engagement measures, we apply principal component analysis (PCA) on these measures to reduce the dimensionality and discover the types of engagement, in line with research question 2**.** Tables 7 and 8 illustrate the PCA results, which align with three broad categories of collective engagement. We label these dimensions as nonengagement, sustained attention driven engagement, and selective attention driven engagement. We define nonengagement as occurring when participants are not engaged either by their cognitive mental models or when they seek information that is not of interest to them. In most theoretical studies relating to engagement, two underlying dimensions are posited to be selective attention and sustained attention (Peters et al. 2009). Selective attention is a basic form of engagement, such as a user quickly glancing at a video that has the potential of relevant interest. Selective attention is required to trigger a point of engagement (O'Brien and Toms 2010). Sustained attention provides a more elaborate form of engagement and allows the possibility of showing affect, feedback, and interaction (Peters et al. 2009). Sustained engagement is marked by participants' attention being maintained in the interactions. This is achieved through the presentation of feedback, and other novel features on the interface.

The first principal component loads negatively with almost all of the collective engagement measures, revealing collective nonengagement with the video. A user searching for a particular medical keyword may be presented with a list of videos. Either the content of the videos may not be appealing to the user, or the user is not enthusiastic about the video showed. The second principal component loads with several measures of views and comments, and so forth, which we label as sustained attention driven engagement. When users have more sustained engagement with the videos, they interact more with the videos by leaving comments, likes, and dislikes. The third principal component loads strongly with measures of relevance, which reveals the extent to which the content of a video matches the interests of users. Therefore, we label this dimension as selective attention driven engagement.

Table 5. Descriptive Statistics of Online Video Collective Engagement Measures

Perspectives	*Measures*	*Min*	*Q_1*	*Median*	*Mean*	*Q_3*	*Max*
Feedback	# of likes	0	6	61.5	5847.8	366.5	1,480,673
	# of dislikes	0	0	2	122.91	14.75	30,529
	# of comments	0	1	8	436	44	80,732
Affect	# of positive comments	0	0	4	13.57	18	91
	# of negative comments	0	0	0	3.7	3.75	44
	# of neutral comments	0	0	1	5.28	5	65
Motivation	sim(keyword, title)*	0	0.17	0.34	0.34	0.48	1
	sim(keyword, description)	0	0	0.07	0.11	0.16	1
Interest	sim(comments, description)	0	0	0.27	1.27	1.53	14.12
	sim(comments, title)	0	0	0.17	0.8	0.98	1
	# of views	0	1,460	12,112	522,659	67,363	1,332,772,312

*sim(*A*, *B*) = cosine(*A*, *B*), the cosine similarity.

Table 6. Correlation of Collective Engagement Measures

	Numeric Variables	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1	# of views	1										
2	# of dislikes	**0.89**	1									
3	# of neutral comments	0.33	0.34	1								
4	sim(comment, description)	0.09	0.11	0.65	1							
5	# of likes	**0.93**	**0.96**	0.33	0.09	1						
6	# of comments	**0.85**	**0.88**	0.39	0.12	0.91	1					
7	sim(keyword, title)	-0.13	-0.12	-0.18	-0.15	-0.13	-0.14	1				
8	sim(keyword, description)	-0.07	-0.07	-0.12	-0.06	-0.08	-0.08	0.22	1			
9	sim(comment, title)	0.18	0.21	**0.70**	**0.77**	0.20	0.20	-0.15	-0.09	1		
10	# of negative comments	0.16	0.18	**0.71**	**0.72**	0.15	0.21	-0.13	-0.09	**0.76**	1	
11	# of positive comments	0.17	0.17	0.71	0.76	0.17	0.17	-0.16	-0.14	0.78	0.66	1

Table 7. Principal Component Analysis

Component	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
Standard Deviation	2.18	1.72	1.08	0.88	0.60	0.56	0.48	0.42	0.38	0.32	0.17
Proportion of Variance	**0.43**	**0.27**	**0.11**	0.07	0.03	0.03	0.02	0.02	0.01	0.01	0.00
Cumulative Proportion	**0.43**	**0.70**	**0.81**	0.88	0.91	0.94	0.96	0.98	0.99	1.00	1.00

Table 8. Selected Components and Loadings

Component & Loadings	*Nonengagement*	*Sustained Attention*	*Selective Attention*
# of views	**-0.325**	**0.369**	0.075
# of dislikes	**-0.335**	**0.368**	-0.089
# of neutral comments	**-0.363**	**-0.203**	-0.017
sim(comment, description)	**-0.299**	**-0.354**	-0.065
# of likes	**-0.335**	**0.383**	0.097
# of comments	**-0.335**	**0.347**	0.086
sim(keyword, title)	0.118	0.091	**0.666**
sim(keyword, description)	-0.061	0.08	**0.730**
sim(comment, title)	**-0.335**	**-0.313**	0.007
# of negative comments	**-0.318**	**-0.308**	-0.045
# of positive comments	**-0.321**	**-0.317**	-0.023

Results and Discussion

Propensity Score Matching and Causal Effect Estimation

We are interested in the relationship between encoded medical information in videos and different dimensions of collective engagement: nonengagement, sustained attention driven engagement, and selective attention driven engagement. A challenge in estimation is the endogeneity between collective engagement and medical information in videos. First, a video's ability to engage with viewers, as well as the medical information contained in the video, could be influenced by external factors, such as the credibility of the channel posting the video. Second, there could also be systematic differences in how users engage with informative vs. versus non-informative videos depending on unobserved taste preferences and channel heterogeneity in attracting viewers. Third, a channel may have a higher incentive to create video materials that provide medical information when they observe greater engagement with the video. The fundamental challenge in identification is that we do not observe the counterfactual; that is, it is not currently possible to randomize treatment across a channel by showing videos with high/low medical information to viewers and measuring the impact of engagement of a video.

Given the difficulty in identifying whether it is the encoded medical information in videos that triggered collective engagement, matching on propensity scores provides a means to identify the actual causal impact. A channel's propensity to post a video high in medical information is considered the treatment. The effect of medical information on collective engagement is estimated as the average treatment effect. Identification is possible since we have (1) a set of video level measures and channel characteristics that are distinct from medical information encoded in the video, (2) a method of classification whereby we can categorize videos as containing a high level of encoded medical information (the treatment condition), and (3) post-treatment collective engagement with the video. Since the measures of collective engagement are available only after the video is posted by the channel, we temporally separate the treatment (the medical information encoded in a video) from the post-treatment outcome, the collective engagement.

We train a classifier with content-related video-level features to classify videos into high medical information videos and low medical information videos. Figure 4 describes the procedure for annotation and classification of encoded medical information as well as our causal identification approach.

We examine videos with a similar propensity in having a high level of encoded medical information and conduct a matched sample analysis. Our matching approach addresses the issue that endogeneity in the relationship between medical information encoded in the video and collective engagement could result from a content creator's ability to retain engaged viewers. Table 9 shows a summary of the logistic regression classifier for video medical information classification. The model is trained on 480 videos and evaluated on 120 videos. The number of unique medical terms is a significant predictor for high medical information encoded in a video.

The performance of the classification is reported in Table 10. We deploy this classification model to classify the remaining videos in our collection. In total, we have 7,948 videos classified as high medical information and 11,325 videos classified as low medical information.

Since collective engagement with a video may be influenced by a whole host of factors external to the content of a video, we control for heuristic measures of video quality such as duration of the video (Pandey et al. 2010; Sood et al. 2011; Steinberg et al. 2010), titles and tags (Figueiredo et al. 2009), good description or a comprehensive narrative (Gooding et al. 2011), technical quality (light, sound, resolution; Gooding et al. 2011; Steinberg et al. 2010), credentials (Gooding et al. 2011), and evidence-based practices used in the video (Dawson et al. 2011). In Table 11, we report the difference in means of the numeric variables of videos in the high and low medical information groups. T-test results show that values for all five numeric variables in the two groups are significantly different.

We model a video's propensity to encode medical information using a logistic regression model with several observable channel- and video-specific factors to account for endogeneity and heterogeneity. Table 12 shows the propensity score estimation results using logistic regression. The dependent variable is the encoded medical information (which is a binary variable denoting high or low medical information).

We then use the estimated propensity scores to conduct pairwise matching between the treatment group of videos that encode a high degree of medical information and the control group of videos that resemble the treatment group of videos, based on similar propensity scores, using a genetic matching method (Diamond and Sekhon 2013). Figure 5 shows the propensity score by treatment status. We conducted several tests to ensure the validity of our propensity score modeling approach. We first ensured that a comparison of histograms reveals there is common support for the treatment and the control groups. We also ensured that our logistic regression accurately predicts membership in the treatment group (i.e., the propensity of a video to encode high medical information).

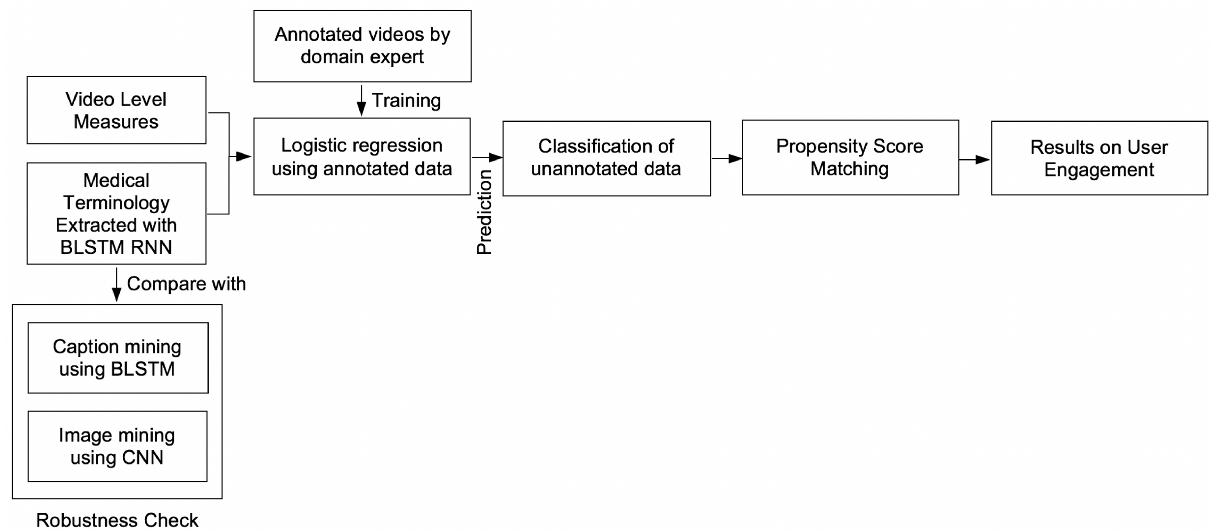

Figure 4. Propensity Score Matching and Robustness Checks

Table 9. Logistic Regression Model Summary

	Estimate	*Std. Error*	*Z value*	*P-value*
(Intercept)	-1.235	0.100	-13.303	< 0.01
# of unique words in description	0.008	0.002	3.784	< 0.01
# of words in description	-0.002	0.001	-2.37	< 0.05
# of unique medical terms	0.007	0.001	11.245	< 0.01
Video definition	-0.765	0.125	-6.164	< 0.01
Video duration(s)	0.001	0.001	-0.472	0.637
Has tags	0.144	0.091	1.578	0.115
Has caption	1.086	0.153	7.099	< 0.01

Table 10. Medical Information Classification Evaluation Results

	Precision	*Recall*	*F-measure*
High Medical Information Videos	88.7%	83.9%	86.2%
Low Medical Information Videos	86.6%	90.6%	88.6%
	Overall Accuracy: 87.5%		

Table 11. Pre-matching Difference-in-Means

Groups	**log(# of channel comments)**	**log(# of channel subscribers)**	**log(# of channel views)**	**# of published days**	**log(avg. view count)**
Mean of low medical information videos	0.80	6.05	11.83	1165.6	6.93
Mean of high medical information videos	1.02	8.36	14.07	811.6	8.33
T-test (*p*-value)	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

Table 12. Propensity Score Estimation

	Estimate	*P*-value
(Intercept)	-1.048	< 0.01
log(# of channel views)	-1.57510^{9}	< 0.01
log(# of channel subscribers)	563.2	< 0.01
log(# of channel comments)	3.061	< 0.01
log(# of channel video count)	-3.042	< 0.01
log(# of channel average video view count)	-2.918	< 0.01
# of published days	-6.83810^{-4}	< 0.01
Content creator credibility	93.6	< 0.01
Video definition (SD)	0.22	< 0.01

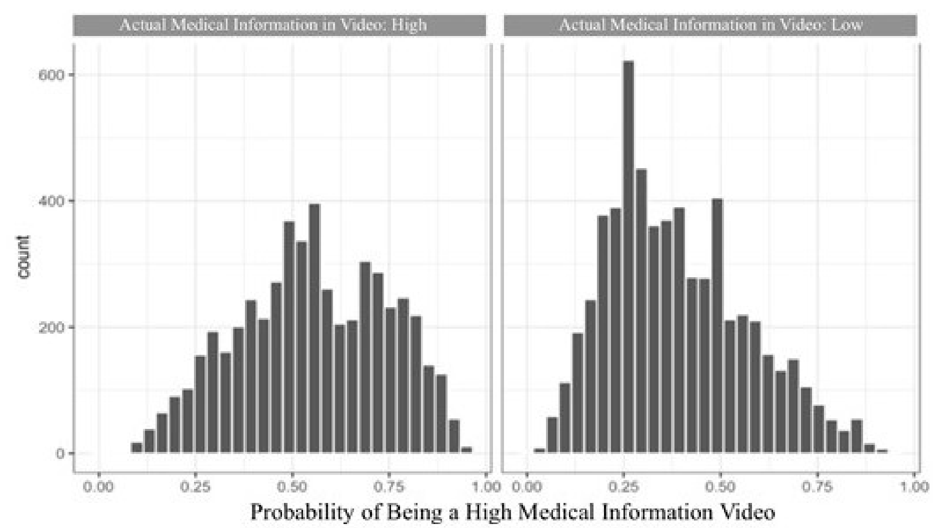

Figure 5. Estimated Propensity Scores by Treatment States

The likelihood that a video encodes high medical information is considered as an endogenous treatment that results from observed channel characteristics. Conditioning on observables, we can ignore treatment into a selection regime (Rosenbaum and Rubin 1984), that is, the fact that there is a systematic difference in posting of high versus low medical information videos. We, therefore, examine videos with a similar propensity to contain highly informative medical content and conduct a matched sample analysis. The results of the matched sample analysis are shown in Table 13. The differences in the numeric variable values are not significant anymore.

Determining whether collective engagement resulted from high medical information or intrinsic video characteristics is a significant challenge. The endogeneity issue in this study is two-fold: (1) whether the medical information in the video is a result of being created by a reputed actor or unobserved motivations of content creators to post highly informative medical videos on YouTube, or (2) whether the collective engagement with the video arises from unobserved user preferences or results from video-specific factors. Matching on propensity scores provides us a way to identify the impact of medical information on collective engagement. Table 14 illustrates the estimated treatment effect on the three principal components of collective engagement.

We find that medical information is significantly negatively linked to the measure of nonengagement, suggesting that users are less likely to engage with videos if they do not contain medical information. Conversely, it also implies that videos that contain valuable medical information are more engaging to a user. Secondly, we find a positive and marginally significant relationship between medical information encoded in the video and selective attention driven engagement. This suggests that high medical information has the potential

Table 13. Post-Matching Difference-in-Means

Groups	log(# of channel comments)	log(# of channel subscribers)	log(# of channel views)	# of published days	log(avg. view count)
Mean in low medical information videos	0.69	6.60	11.96955	787.73	6.613416
Mean in high medical information videos	0.82	7.10	12.59441	671.28	6.770645
T-test (*p*-value)	0.78	0.65	0.61	0.59	0.85

Table 14. Estimated Treatment Effect

Component		*Estimate*	*P-value**
Nonengagement	(Intercept)	0.78	< 0.01
	Average Treatment Effect	-2.98	< 0.01
Sustained attention	(Intercept)	0.48	< 0.01
	Average Treatment Effect	-1.54	< 0.01
Selective attention	(Intercept)	-0.05	0.18
	Average Treatment Effect	0.03	0.09

*P-values are generated with one-tailed test.

to trigger engagement. Since we consider collective engagement by aggregate groups of users, it is possible that different users have different informational needs and some groups of users are engaged by more medical information. At the same time, since we also find a negative relationship between medical information and sustained attention driven engagement, our results suggest that healthcare professionals need to have a nuanced understanding of what drives engagement. It is possible that users engage with medical information in a superficial manner (resulting in the positive relationship between medical information and selective attention driven engagement) while the complexity of encoded medical information may discourage users from deeper interaction due to the high cognitive load to comprehend the information (resulting in the negative relationship between medical information and sustained attention driven engagement).

Robustness Checks

Our method of extracting medical terminology from video descriptions and using it to assist the classification of medical information in the videos is useful but not exhaustive, as medical information is also conveyed in a video's narrative and content. Our proposed approach relies on video description for the following reasons: First, video descriptions are usually contributed by a content creator when they submit the videos to the platform. Second, only 11% of videos contain downloadable closed captions submitted by content creators. The automatic captions are inaccurate and unavailable publicly. It is not feasible to consistently obtain video captions for information extraction and classification. Third, YouTube video content is not publicly available. Information extraction from the video also requires significant annotation effort. As a robustness check, we compare our proposed approach to classifying medical information with additional medical terms extracted from closed captions and video content.

Medical Information Classification with Video Captioning

We apply the BLSTM RNN model to the closed captions from 600 annotated videos to extract medical terms from the captions. We construct a logistic regression model for medical information classification with the number of unique medical terms from closed captions instead of video descriptions. We examine the impact of information extraction from caption on the video medical information classification performance. The performance of the classification is reported in Table 15. Overall, using medical terminology extracted from closed captions increases classification accuracy and recall of high medical information videos and precision of low medical information videos. However, the improvement in overall accuracy is not statistically significant on the test data (*p*-value = 0.322).

Robustness Checks for Medical Information Classification with Image Analytics

Image classification performance has been significantly improved with the advent of the large datasets such as ImageNet

Table 15. Medical Information Classification with Video Captioning Evaluation Results

	Precision	*Recall*	*F-measure*
High Medical Information Videos	86.2%	89.3%	87.7%
Low Medical Information Videos	90.3%	87.5%	88.9%
Overall Accuracy	88.3%		

Table 16. Medical Information Classification with Image Analytics Evaluation Results

	Precision	*Recall*	*F-measure*
High Medical Information Videos	86.5%	80.5%	83.4%
Low Medical Information Videos	83.8%	84.3%	84.0%
Overall Accuracy	85.0%		

using convolutional neural network (CNN) architectures such as AlexNet, Inception, and ResNet. We examine whether CNN yields good performance on image classification in video frames and if these class labels enhance medical information classification performance. We construct a large video frame dataset by sampling one frame every two seconds from the 600 annotated videos. In total we extracted 204,713 frames. We define our task as predicting the objects in each frame. We utilize an Inception CNN architecture that was pre-trained on approximately 1.28 million images (1,000 object categories) from the 2014 ImageNet Large Scale Visual Recognition Challenge (Szegedy et al. 2016). Image classification is performed a workstation with 32G RAM and NVIDIA Quadro K420 GPU. A batch of 100 frames requires 320 seconds to process. We predict the probability of the 1,000 object categories for each frame using the inception CNN architecture and extract the top five object categories for each frame. Object category labels with low probability are not considered for further analysis. We identify the number of medically relevant objects at the video level. Moreover, we incorporate the number of medically relevant objects as an additional feature in the logistic regression for medical information classification. The performance of the classification is reported in Table 16.

We find that incorporating object categories negatively affects the performance of our classification method. This could be because the Inception CNN model is not trained with diabetes video frames but with ImageNet, and the object categories are too general and contain noise that may affect the classification performance. To improve object detection in healthcare videos, future work may need to include methods such as transfer learning within the Inception CNN model. Significant effort will be needed to develop a set of medical relevant objects for chronic care and annotate the video frames for transfer learning. Another issue for future work is to develop a set of ontologies specific to chronic care.

Implications for Literature and Future Research

This paper contributes to three distinct streams of research. The first stream of research involves the applications of recent advances in deep learning to healthcare informatics. Our study is among the first to utilize a scalable deep learning approach to assessing complex medical information from real-world user-generated video data. We examine different approaches to extract medical information from videos: information extraction from text (i.e., video descriptions and closed captions) and object detection from video frames. Deep learning methods have shown great promise in information extraction, natural language understanding, and image classification, especially in an era of patient-centric care and precision medicine (Miotto et al. 2017). Applications of novel deep learning methods contribute to superior performance as compared to conventional machine learning methods and consequently to the robustness and rigor of our research. These applications may open new opportunities for data-driven research with electronic medical records that include text, video and images, social media, mobile, and sensor data.

The second stream of work is in understanding user-generated content literature in IS research. We combine machine learning with statistical and econometric methods to evaluate the impact of medical information on collective engagement on video sharing platforms. Videos with a high degree of medical information may intimidate the users from commenting on and liking the content, particularly given the challenges in health literacy. Our approach goes beyond scraping data from YouTube to provide a systematic, evidence-based, transparent and generalizable method to retrieve videos. Our approach can be extended to understand retrieval of information in user-generated content platforms, and more broadly to understand how content triggers and maintains a sustained engagement from users in other domains. This approach can also be aplied to understand how brands should manage their social media content to better engage and communicate with their customers.

The third stream of research is patient education and communication afforded by technology and online solutions defined by the new area of digital therapeutics. The Internet has reduced much of the information asymmetry between healthcare practitioners and consumers by providing multiple avenues whereby patients can educate themselves with user-generated content. A better understanding of how patients engage with medical information offers policy implications for the utilization of healthcare resources and the quality of delivered care. While healthcare organizations or consumers produce patient educational materials, not only should they think about what the medical information to deliver but also how to meet the interest, information needs and health literacy level of the consumers. Overwhelming patients with complex medical information could be counterproductive for patient education. Digital therapeutics that uses technology and online solutions to reach patients with complex chronic conditions with personalized, contextualized and just-in-time content is necessary for improving health literacy and outcomes. Identifying and recommending YouTube videos, with high medical information encoded, that potentially meet this need is a key contribution of this stream of research.

We synthesize multiple machine learning methods with a process of annotation by domain experts to design a causal framework that permits us to automate the task of extracting medical information and learn group behavior in watching health related videos. Furthermore, the nature of the data, which does not include individual level usage info but aggregate usage data, results in new insights about collective attention. The future research directions from our study are twofold. One is that analysis of aggregate video usage data in a variety of contexts can be used to generate collective level insights. Second, our study provides a foundation for future work to examine the relationship between content quality and information encoded in content at a more granular level with individual level video usage data.

Implications for Practice and Chronic Care

As explained earlier, prior approaches for assessing medical information in videos relied heavily on the judgment of domain experts to evaluate the content (Backinger et al. 2011; Dawson et al. 2011). As the volume of online videos grows exponentially, using expert evaluation to assess the videos on YouTube is not a sustainable long-term solution. We design and operationalize a scalable approach based on deep learning, thereby overcoming the inability of expert-based assessments annotating medical information in large-scale datasets of user created videos. Our findings highlight gaps in how aggregate groups of users perceive and engage with medical information.

According to the chronic care model, chronic condition management can only be achieved through a multifaceted approach encompassing patient support groups and partnerships between patient communities and healthcare professionals (Coleman et al. 2009). Traditionally, patient education has been viewed as a physician-centered initiative (Jordan et al. 2008). It is now recognized that both patients and physicians share the responsibility in healthcare decision-making. Healthcare professionals need to provide decision support and patient self-management support based on evidence and patients' needs (Warm 2007). It is recognized that a coordinated and planned approach is necessary so that engaged patients utilize available medical and community resources (Powers et al. 2017). Given the prohibitive cost of preparing educational material, clinicians can incorporate high quality user-generated videos into clinical practice and use them as a digital therapeutic interventions similar to drug therapy. Our method can be used to identify appropriate videos for self-management education and decision support across the continuum of disease progression and care delivery, from diagnosis of the chronic condition to final stages of the disease, as explained in Table 17. Our methods can also be applied to engage other stakeholders, such as caregivers and family members, to utilize user-generated videos in the chronic disease management process.

We find a tradeoff in collective engagement depending on the level of medical information encoded in videos. This implies that well-informed patients are likely to engage with high levels of medical information, while patients with low levels of information are nonengaged. This raises the challenges with visual social media as an enabler of patient education. Social media could propagate misinformation for patients with chronic conditions. Further, uninformed patients find it difficult to comprehend complex medical information for decision-making and self-management, leading to nonengagement when they encounter videos high in medical information. Our results suggest the following healthcare guidelines for patient education and collective engagement. Our method can be used to develop both best practice guidelines and a method of automated video retrieval that could work across the spectrum of patient engagement and levels of understanding of medical information (Table 18).

Limitations and Directions for Future Research

Our study has several areas for future improvements. First, our image analysis utilizes a pre-trained Inception model with the ImageNet dataset. The ImageNet dataset contains 1,000 classes

Table 17. Implications for Chronic Care

Element of Chronic Care	*Goals Highlighted by the Literature*	*The Contribution of Our Approach*
Patient Awareness	Information of the condition, awareness of signs and symptoms	Identifying and retrieving videos with the most sustained attention driven engagement
Self-management support	Managing symptoms, drug interactions, adherence to treatment regimen	Identifying and retrieving videos that engage patients in treatment compliance
Disease management for high-risk patients	Treatment and monitoring	Designing videos for sustained attention in enabling self-monitoring and interactions with the providers

Table 18. Implications of Our Work for Content Creators and Healthcare Practitioners

	Collective Engagement: Low	Collective Engagement: High
Medical Information: Low	Our method enables an automated video retrieval approach to identify and label misleading videos	Our method provides a scalable approach for automated retrieval of high quality user generated content that enables both selective and sustained engagement
Medical Information: High	Healthcare organizations can use our method to develop guidelines for engagement with complex medical information	Our work can be used to develop best practices for healthcare organizations to promote sustained engagement

of general objects. The classification results are not tailored to the medical context. Constructing a labeled dataset to identify medical related objects in the image requires significant time and research effort. We will need a controlled vocabulary of medical objects and annotation of a significant amount of images guided by domain experts, which is beyond the scope of our current study. In the future, we plan to explore transfer learning for chronic care video understanding with necessary annotations.

Second, our proposed research approach mainly relies on text descriptions in YouTube video metadata to classify encoded medical information in the videos. We perform robustness checks by examine the possibility of using video closed captions and video frames to enhance the performance of encoded medical information classification. However, our results suggest that incorporating closed captions and video frames do not improve the classification of medical information. Notably, closed captions are only available for 11% of the videos in our collection. This distribution may reflect the sparsity in closed captions among all the videos on YouTube. Current speech-to-text technology generates automatic transcripts for videos but does not guarantee a satisfying level of accuracy. Off-the-shelf convolutional neural network architectures without proper training and annotation do not provide accurate and relevant image classification results. Future work could focus on developing an integrated learning framework for videos related to chronic care leveraging three different sources of data: video metadata, captions, and frames. Sophisticated deep learning models could improve the performance of automatic transcription generation and medical object detection.

Another limitation of our work is that we develop a set of measures for collective engagement that are reliant on YouTube metadata. Since we employ PCA, the three measures of engagement are orthogonal by construction. Future work may identify engagement across the continuum. Some measures of engagement considered in prior literature on user engagement such as focused attention, control, aesthetics, and so on, are not observable from YouTube metadata. In the future, we would like to conduct a controlled experiment to acquire more comprehensive measures of collective engagement on videos. An added dimension may be the complexity and depth of medical information in the videos. Currently, we categorize the information broadly as just medical information, but this could be at either a superficial level or in depth, which also influences its reception and engagement by different categories of users. The impact of prior health literacy of a user on these categories of engagement is one that can be studied in a future experimental setting. A two dimensional exploration of individual level engagement with videos and granular level medical informational content in videos will enable a better matching of videos to users to assess patient engagement at scale.

Fourth, while medical information can be empowering consumers to make important health decisions, it can also be confusing and overwhelming. While the wealth of medical information available on the Internet could be harnessed for patient education and empowerment, the body of available medical knowledge is continually being updated as a result of new research. With rapid advances in natural language understanding and video understanding, newer methods in text and video mining need to be brought together with state of art medical knowledge to assess the contextualized quality of medical information disseminated on social media platforms.

Conclusions

In this study, we synthesize deep learning and econometric methods to understand collective engagement with health information, specifically YouTube videos on diabetes. We first extract medical terminology with a deep learning architecture, BLSTM RNN , and classify medical information encoded in videos from YouTube using a logistic regression with medical terms and other video level features. We then examine how collective engagement varies with the medical information encoded in videos. Our findings have profound implications for information systems researchers, healthcare professionals, patient educators, and policymakers. Based on our findings, we also propose normative guidelines for content creators and healthcare practitioners to produce engaging and relevant patient education materials. Our study contributes to chronic care management by better connecting patients and caregivers with community resources and providing patient-centered self-management and decision-making support with digital therapeutics.

Acknowledgments

We thank participants at the Workshop for Information Systems and Technology, Workshop on Health IT and Economics, the China Summer Workshop on Information Management, the Workshop on Information Systems Economics, Statistical Challenges in Ecommerce Research, and workshop participants at the University of Texas at Austin, Georgia State University, and Arizona State University for comments on earlier versions of this paper. We also thank the senior editors, associate editor, and reviewers as well as participants at the *MIS Quarterly* Special Issue Workshop for their comments.

References

Adams, R. J. 2010. "Improving Health Outcomes with Better Patient Understanding and Education," *Risk Management and Healthcare Policy* (6:31), pp. 61-72.

Attfield, S., Kazai, G., Lalmas, M., and Piwowarski, B. 2011. "Towards a Science of User Engagement (Position Paper)," in *WSDM Workshop on User Modeling for Web Applications*, Hong Kong, February 9-12 (http://www.dcs.gla.ac.uk/~mounia/Papers/engagement.pdf)

Asur, S., Huberman, B. A., Szabo, G., and Wang, C. 2011. "Trends in Social Media: Persistence and Decay," in *Proceedings of Fifth International AAAI Conference on Weblogs and Social Media*, Barcelona, July 17-21, pp. 434-437.

Backinger., C. L., Pilsner, A. M., Augustson, E. M., Frydl, A., Phillips, T., and Rowden, J. 2011. "YouTube as a Source of Quitting Smoking Information," *Tobacco Control* (20:2), pp. 119-122.

Baldwin, T., Kim, Y. B., De Marneffe, M. C., Ritter, A., Han, B., and Xu, W. 2015. "Shared Tasks of the 2015 Workshop on Noisy User-Generated Text: Twitter Lexical Normalization and Named Entity Recognition," *Proceedings of the ACL 2015 Workshop on Noisy User-Generated Text*, Beijing, July 31, pp. 126-135.

Bauer, U. E., Briss, P.A., Goodman, R. A., and Bowman, B. A. 2014. "Prevention of Chronic Disease in the 21st Century: Elimination of the Leading Preventable Causes of Premature Death and Disability in the USA," *The Lancet* (384:9937), pp. 45-52.

Bengio, Y., Simard, P., and Frasconi, P. 2014. "Learning Long-Term Dependencies with Gradient Descent Is Difficult," *IEEE Transactions on Neural Networks* (5:2), pp. 157-166.

Bodenreider, O. 2004. "The Unified Medical Language System (UMLS): Integrating Biomedical Terminology," *Nucleic Acids Research* (32:1), pp. 267-270.

Carman, K. L., Dardess, P., Maurer, M., Sofaer, S., Adams, K., Bechtel, C., and Sweeney, J. 2013. "Patient and Family Engagement: A Framework for Understanding the Elements and Developing Interventions and Policies," *Health Affairs* (32:2), pp. 223-231.

CDC. 2017. *National Diabetes Statistics Report, 2017*, Atlanta, GA: Centers for Disease Control and Prevention, U.S. Department of Health and Human Services.

Chatzopoulou, G., Sheng, C., and Faloutsos, M. 2010. "A First Step Towards Understanding Popularity in YouTube," in *Proceedings of the INFOCOM IEEE Conference on Computer Communications Workshops*, New York: IEEE, pp. 1-6.

Chou, W. Y. S., Prestin, A., Lyons, C., and Wen, K. Y. 2013. "Web 2.0 for Health Promotion: Reviewing the Current Evidence," *American Journal of Public Health* (103:1), pp. e9-e18.

Coleman, K., Austin, B. T., Brach, C., and Wagner, E. H. 2009. "Evidence on the Chronic Care Model in the New Millennium," *Health Affairs* (28:1), pp. 75-85.

Crane, R., and Sornette, D. 2008. "Robust Dynamic Classes Revealed by Measuring the Response Function of a Social System," *Proceedings of the National Academy of Sciences* (105:41), pp. 15649-15653.

Dawson, A. L., Hamstra, A. A., Huff, L. S., Gamble, R. G., Howe, W., Kane, I., and Dellavalle, R. P. 2011. "Online Videos to Promote Sun Safety: Results of a Contest," *Dermatology Reports* (3:1).

Diakopoulos, N. A. 2015. "The Editor's Eye: Curation and Comment Relevance on *The New York Times*," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work and Social Computing*, New York: ACM, pp. 1153-1157.

Diamond, A., and Sekhon, J. S. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies," *Review of Economics and Statistics* (95:3), pp. 932-945.

Dos Santos, C. D., and Zadrozny, B. 2014. "Learning Character-Level Representations for Part-of-Speech Tagging," in *Proceedings of the 31st International Conference on Machine Learning*, Beijing, pp. 1818-1826.

Eysenbach, G., and Jadad, A. R. 2001. "Evidence-Based Patient Choice and Consumer Health Informatics in the Internet Age," *Journal of Medical Internet Research* (3:2), p. e19.

Eysenbach, G., and Kohler, C. 2002. "How Do Consumers Search for and Appraise Health Information on the World Wide Web? Qualitative Study Using Focus Groups, Usability Tests, and In-Depth Interviews," *British Medical Journal* (324:7337), pp. 573-577.

Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., and Stern, A. 2004. "Health Related Virtual Communities and Electronic Support Groups: Systematic Review of the Effects of Online Peer to Peer Interactions," *British Medical Journal* (328:1166).

Eysenbach, G., Sa, E. R., and Diepgen, T. L. 1999. "Shopping Around the Internet Today and Tomorrow: Towards the Millennium of Cybermedicine," *British Medical Journal* (319:7220).

Fernandez-Llatas, C., Traver, V., Borras-Morell, J. E., Martinez-Millana, A., and Karlsen, R. 2017. "Are Health Videos from Hospitals, Health Organizations, and Active Users Available to Health Consumers? An Analysis of Diabetes Health Video Ranking in YouTube," *Computational and Mathematical Methods in Medicine* (8194940).

Figueiredo, F., Belém, F., Pinto, H., Almeida, J., Gonçalves, M., Fernandes, D., and Cristo, M. 2009. "Evidence of Quality of Textual Features on the Web 2.0," in *Proceedings of the 18th ACM Conference on Information and Knowledge Management*, New York: ACM, pp. 909-918.

Fleckenstein, K. S. 1991. "Defining Affect in Relation to Cognition: A Response to Susan McLeod," *Journal of Advanced Composition* (11:2), pp. 447-453.

Fox, S., and Purcell, K. 2010. "Chronic Disease and the Internet," Pew Research Center, Washington, DC.

Gooding, L. F., and Gregory, D. 2011. "Descriptive Analysis of YouTube Music Therapy Videos," *Journal of Music Therapy* (48:3), pp. 357-369.

Heathfield, H., and Louw, G. 1999. "New Challenges for Clinical Informatics: Knowledge Management Tools," *Health Informatics Journal* (5:2), pp. 67-73.

Hernandez-Tejada, M. A., Campbell, J. A., Walker, R. J., Smalls, B. L., Davis, K. S., and Egede, L. E. 2012. "Diabetes Empowerment, Medication Adherence and Self-care Behaviors in Adults with Type 2 Diabetes," *Diabetes Technology & Therapeutics* (14:7), pp. 630-634.

Hersh, W. R. 2002. "Medical Informatics: Improving Healthcare Through Information," *JAMA* (288:16), pp. 1955-1958.

Hochreiter, S., and Schmidhuber, J. 1997. "Long Short-Term Memory," *Neural Computation* (9:8), pp. 1735-1780.

Hutto, C. J., and Gilbert, E. 2014. "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," in *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media*.

Institute of Medicine. 2004. *Health Literacy: A Prescription to End Confusion,* Washington, DC: Institute of Medicine, Board on Neuroscience and Behavioral Health, Committee on Health Literacy.

Jordan, J. E., Briggs, A. M., Brand, C. M., and Osborne R. H. 2008. "Enhancing Patient Engagement in Chronic Disease Self-management Support Initiatives in Australia: The Need for an Integrated Approach," *Medical Journal of Australia* (189:10), pp. S9-S13.

Kallinikos J., and Tempini, N. 2014. "Patient Data as Medical Facts: Social Media Practices as a Foundation for Medical Knowledge Creation," *Information Systems Research* (25:4) pp. 817-833.

Keselman, A., Browne, A. C., and Kaufman, D. R. 2008. "Consumer Health Information Seeking as Hypothesis Testing," *Journal of the American Medical Informatics Association* (15:4), pp. 484-495.

Keelan, J., Pavri-Garcia, V., Tomlinson, G., and Wilson, K. 2007. "YouTube as a Source of Information on Immunization: A Content Analysis," *JAMA* (298:21), pp. 2482-2484.

Kutner, M., Greenburg, E., Jin, Y., and Paulsen, C. (2006. "The Health Literacy of America's Adults: Results from the 2003 National Assessment of Adult Literacy," NCES 2006-483, National Center for Education Statistics, U.S. Department of Education, Washington, DC.

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., and Dyer, C. 2016. "Neural Architectures for Named Entity Recognition," arXiv preprint arXiv:1603.01360.

Lau, A. Y., Gabarron, E., Fernandez-Luque, L., and Armayones, M. 2012. "Social Media in Health—What Are the Safety Concerns for Health Consumers?," *Health Information Management Journal* (41:2), pp. 30-35.

Laurel, B. 1993. *Computers as Theatre*, Reading, MA: Addison-Wesley.

LeCun Y., Bengio, Y., and Hinton, G. 2015. "Deep Learning," *Nature* (521), pp. 436-444.

Mackert, M., Mabry-Flynn, A., Champlin, S., Donovan, E. E., and Pounders, K. 2016. "Health Literacy and Health Information Technology Adoption: The Potential for a New Digital Divide," *Journal of Medical Internet Research* (18:10), p. e264.

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., and Gramopadhye, A. K. 2015. "Healthcare Information on YouTube: A Systematic Review," *Health Informatics Journal* (21:3), pp. 173-194.

Mao, J. J., Chung, A., Benton, A., Hill, S., Ungar, L., Leonard, C. E., and Holmes, J. H. 2013. "Online Discussion of Drug Side Effects and Discontinuation among Breast Cancer Survivors," *Pharmacoepidemiology and Drug Safety* (22:3), pp. 256-262.

Milokov, T., Sutskever, I., Chen, K., Corrando, G. S., and Dean, J. 2013. "Distributed Representations of Words and Phrases and Their Compositionality," *Advances in Neural Information Processing Systems 26* (https://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf).

Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. 2017. "Deep Learning for Healthcare: Review, Opportunities and Challenges," *Briefings in Bioinformatics* (19:6), pp. 1236-1246.

Moorhead, S. A., Hazlett, D. E., Harrison, L., Carroll, J. K., Irwin, A., and Hoving C. 2013. "A New Dimension of Health Care: Systematic Review of the Uses, Benefits, and Limitations of Social Media for Health Communication," *Journal of Medical Internet Research* (15:4).

Munezero, M. D., Montero, C. S., Sutinen, E., and Pajunen, J. 2014. "Are They Different? Affect, Feeling, Emotion, Sentiment, and Opinion Detection in Text," *IEEE Transactions on Affective Computing* (5:2), pp. 101-111.

Napoli, P. 2011. *Audience Evolution: New Technologies and the Transformation of Media Audiences*, New York: Columbia University Press.

O'Brien, H. L., and Toms, E. G. 2008. "What Is User Engagement? A Conceptual Framework for Defining User Engagement with Technology," *Journal of the Association for Information Science and Technology* (59:6), pp. 938-955.

O'Brien, H. L., and Toms, E. G. 2010. "The Development and Evaluation of a Survey to Measure User Engagement," *Journal of the American Society for Information Science and Technology* (61:1), pp. 50-69.

Oh, J., Bellur, S., and Sundar, S. S. 2010. "A Conceptual Model of User Engagement with Media," paper presented at the 60th Annual Conference of the International Communication Association, Singapore.

Pandey, A., Patni, N., Singh, M., Sood, A., and Singh, G. 2010. "YouTube as a Source of Information on the H1N1 Influenza Pandemicm" *American Journal of Preventive Medicine* (38:3), pp. e1-e3.

Pant, S., Deshmukh, A., Murugiah, K., Kumar, G., Sachdeva, R., and Mehta, J. L. 2012. "Assessing the Credibility of the 'YouTube Approach' to Health Information on Acute Myocardial Infarction," *Clinical Cardiology* (35:5), pp. 281-285.

Parker, S., Zipursky, J., Ma, H., Baumblatt, G., Siegel, C. 2018. "A Web-Based Multimedia Program Before Colonoscopy Increased Knowledge and Decreased Anxiety, Sedation Requirement, and Procedure Time," *Journal of Clinical Gastroenterology* (52:6), pp. 519-523.

Peck, R. S., Zhou, L. Y., Anthony, V. B., and Madhukar, K. 2008. "Consumer Internet, Bear Stearns Equity Research Report," Bear Stearns, New York.

Peters, C., Castellano, G., and de Freitas, S. 2009. "An Exploration of User Engagement in HCI," in In *Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots*, Boston, November 6, Article 9.

Powers, B., Trinh, J. V., and Bosworth, H. B. 2010. "Can This Patient Read and Understand Written Health Information?," *JAMA* (304:1), pp. 76-84.

Powers, M. A., Bardsley, J., Cypress, M., Duker, P., Funnell, M. M., Fischl, A. H., and Vivian, E. 2017. "Diabetes Self-Management Education and Support in Type 2 Diabetes: A Joint Position Statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics," *The Diabetes Educator* (43:1), pp. 40-53.

Riegel, B., Jaarsma, T., and Strömberg, A. 2012. "A Middle-Range Theory of Self-Care of Chronic Illness," *Advances in Nursing Science* (35:3), pp. 194-204.

Rosenbaum, P. R., and Rubin, D. B. 1984. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *Journal of the American statistical Association* (79:387), pp. 516-24.

Rubenstein, H., Goodenough J. B. 1965. "Contextual Correlates of Synonymy," *Communications of the ACM* (8:10), pp. 627-633.

Ruppert, L., Koster, B., Siegert, A. M., Cop, C., Boyers, L., Karimkhani, C., Winston, H., Mounessa, J., Delavalle, R. P., Reinau, D., Diepgen, T., and Surber, C. 2017. "YouTube as a Source of Health Information: Analysis of Sun Protection and Skin Cancer Prevention Related Issues," *Dermatology Online Journal* (23:1), pp. 1-10.

Rutten, L. J. F., Hesse, B. W., Stauver, J., Wilson, P., Chawla, N., Hartigan, D. B., Moser, R. P., Taplin, S., Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., and Rich, C. 2005. "Explorations in Engagement for Humans and Robots. Artificial Intelligence, (166:1-2), pp. 140-164.

Siersdorfer, S., Chelaru, S., Nejdl, W., and San Pedro, J. 2010. "How Useful Are Your Comments? Analyzing and Predicting YouTube Comments and Comment Ratings," in *Proceedings of the 19th International Conference on World Wide Web*, New York: ACM, pp. 891-900.

Simmons L. E., Wolever, A. M., Bechard, E. M., and Snyderman, R. 2014. "Patient Engagement as a Risk Factor in Personalized Healthcare: A Systematic Review of the Literature on Chronic Disease," *Genome Medicine* (6:2).

Sood, A., Sarangi, S., Pandey, A., and Murugiah, K. 2011. "YouTube as a Source of Information on Kidney Stone Disease," *Urology* (77:3), pp. 558-562.

Steinberg, P. L., Wason, S., Stern, J. M., Deters, L., Kowal, B., and Seigne, J. 2010. "YouTube as Source of Prostate Cancer Information," *Urology* (75:3), pp. 619-622.

Stellefson, M., Chaney, B., Ochipa, K., Chaney, D., Haider, Z., Hanik, B., Chavarria, E., and Bernhadt, J. M. 2014. "YouTube as a Source of Chronic Obstructive Pulmonary Disease Patient Education: A Social Media Content Analysis, *Chronic Respiratory Disease* (11:2), pp. 61-71.

Sutskever, I., Vinyals, O., and Le, Q. V. 2014. "Sequence to Sequence Learning with Neural Networks," *Advances in Neural Information Processing Systems* (4).

Syed-Abdul, S., Fernandez-Luque, L., Jian, W. S., Li, Y. C., Crain, S., Hsu, M. H., and Liou, D. M. 2013. "Misleading Health-Related Information Promoted Through Video-Based Social Media: Anorexia on YouTube," *Journal of Medical Internet Research* (15:2).

Szabo, G., and Huberman, B. A. 2010. "Predicting the Popularity of Online Content," *Communications of the ACM* (53:8), pp. 80-88.

Szegedy, C., Vanhoucke V., Ioffe, S., Shlens, J., and Wojna, Z. 2016. "Rethinking the Inception Architecture for Computer Vision," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826.

Warm, E. J. 2007. "Diabetes and the Chronic Care Model: A Review," *Current Ciabetes Reviews* (3:4), pp. 219-225.

Williams, M. V., Davis, T., Parker, R. M., and Weiss, B. D. 2002. "The Role of Health Literacy in Patient–Physician Communication," *Family Medicine* (34:5), pp. 383-389.

Winker, M. A., Flanagin, A., Chi-Lum, B., White, J., Andrews, K., Kennett, R. L., and Musacchio, R. A. 2000. "Guidelines for Medical and Health Information Sites on the Internet—Principles Governing AMA Web Sites," *JAMA* (283:12), pp. 1600-1606.

Wu, F., and Huberman, B. A. 2007. "Novelty and Collective Attention," *Proceedings of the National Academy of Sciences* (104:45), pp. 17599-17601.

Wyatt, J. C. 2001. "Management of Explicit and Tacit Knowledge," *Journal of the Royal Society of Medicine* (94:1), pp. 6-9.

About the Authors

Xiao Liu is an assistant professor in the Department of Information Systems at Arizona State University. Xiao received her Ph.D. in management information systems from the Eller College of Management at the University of Arizona. She works on health analytics, drug safety surveillance, and social media analytics.

Bin Zhang is an assistant professor at the University of Arizona. He received his Ph.D. from Carnegie Mellon University. He specializes in large social network analysis and statistical modeling of social network problems. His work also focuses on social media, technology diffusion, and business analytics.

Anjana Susarla is an associate professor of Information Systems at the Eli Broad College of Business, Michigan State University. Her research interests include the economics of information systems, contracts and sourcing and artificial intelligence. Her work has appeared in several academic journals and peer-reviewed conferences such as such as *Information Systems Research*, *Journal of Management Information Systems*, *Management Science*, *MIS Quarterly*, and the International Conference in Information Systems, among others. She has served on several editorial boards and has been a recipient of multiple best paper awards. She also publishes op-eds and her work has been quoted in several media outlets.

Rema Padman is the Trustees Professor of Management Science and Healthcare Informatics in the Heinz College of Information Systems and Public Policy at Carnegie Mellon University, and an adjunct professor in the Department of Biomedical Informatics at the University of Pittsburgh School of Medicine. Her research investigates healthcare analytics and operations and data-driven decision support, in the context of clinical and consumer-facing IT interventions in healthcare delivery, such as e-health, m-health, chronic and infectious disease management, and workflow analysis. She has published extensively on healthcare informatics, management science, and information systems research topics.

Appendix A

Video Search Keywords

Diabetes	diabetes supplements	insulin resistance	HemoglobinA1c	insulin infusion pump	Toujeo
diabetes causes	diabetes type	insulin resistance syndrome	high glucose	insulin injection	Tradjenta
diabetes complication	diabetes and depression	insulin secretion	iv glucose tolerance test	insulin needles	Tresiba
diabetes cure	diabetes impotence	insulin sensitive	lifescan	insulin pen injector	Welchol
diabetes depression	about diabetes	islet amyloid polypeptide	normal blood glucose levels	insulin pump	artificial sweetener
diabetes diagnosis	diabetes recipes	Islet cell antibodies	normal glucose tolerance	insulin syringe	diabetes diet
diabetes exercise	diabetes lifestyle	islets of Langerhans	oral glucose challenge	insulin therapy	diabetes prevention
diabetes high risk	diabetes medication side effects	lipodystrophy	Oral glucose tolerance test	intermediate acting insulin	diabetes prevention program
diabetes information	abnormal fat metabolism	obesity	plasma concentration	Invokana	diabetes target range
diabetes medication	abnormal glucose homeostasis	pancreas	plasma glucose	islet cell transplantation	diabetic diet
Diabetes Mellitus	abnormal insulin	pancreatic exocrine disease	post glucose tolerance test	Januvia	glycemic index
diabetes self care	abnormal insulin secretion	pancreatic islet	post prandial glucose	Lantus	physical activity
diabetes self management	adipose tissue	reduced insulin	Post-prandial glucose	Levemir	adult blindness
diabetes supplies	adult-onset diabetes	syndrome X	PPG	linagliptin	Autonomic Neuropathy
diabetes surgery	amylin	Acanthosis nigricans	urine glucose	Lispro	cellulitis
diabetes symptoms	B Cell	body mass index	Acarbose	long acting insulin	diabetes and anorexia
diabetes testing	beta cell failure	body weight	ACTOS	meglitinides	diabetes and female sexual dysfunction
diabetes treatment	beta cell of pancreatic islets	diabetes risk factor	alogliptin	metformin	diabetes and nausea
diabetes type 1 symptoms	diabetes genetic predisposition	diabetes risk factors	alpha-glucosidase inhibitors	miglitol	diabetes and reduced sexual desire women
diabetes type 2 symptoms	diabetes genetics	glycosuria	Apidra	Nateglinide	diabetes and vomiting
diets for diabetes	glucagon	hirsutism	artificial pancreas	Nesina	diabetes infections
DM	glucagon like peptide	hyperandrogenism	Aspart	Novolin	diabetes myocardial infarct

Gestational diabetes	Glucose	increased urination	Avandia	Novolin N	diabetes staph infection
gestational DM	glucose intolerance	increased water intake	bariatric surgery	Novolog	diabetes urinary tract infections
increased risk of diabetes	glucose metabolism	intermediate hyperglycemia	Biguanides	NPH insulin	diabetes UTI
insulin dependent diabetes mellitus	glucose tolerance	PCOS	bolus insulin	Onglyza	diabetic coma
juvenile diabetes	glucose transport	polycystic ovarian syndrome	Canagliflozin	oral hypoglycemic agents	diabetic dermopathy
Type 1	glucose uptake	polydipsia	carbohydrate counting	pancreatic transplantation	diabetic foot
Type 1 diabetes	glycemic control	polyuria	colesevelam	parenteral glucagon emergency	diabetic heart disease
type 1 DM	glycogenolysis	weight loss	Cycloset	pioglitazone	diabetic keto acidosis
Type 2	hepatic gluconeogenesis	weight gain	Detemir	Pramlintide	diabetic mononeuropathy
Type 2 diabetes	HLA complex	blood glucose	diabetes clinical trails	prandin	diabetic nephropathy
Type 2 DM	hyperglycemia	blood pressure	Diabetes medicine	precose	diabetic neuropathy
type I diabetes	hyperinsulinemia	continuous glucose monitoring systems	DPP-4 Inhibitors	rapid acting insulin	diabetic peripheral vascular disease
type I DM	IAPP	c-peptide test	FlexPen	Regular insulin	diabetic polyneuropathy
Type II	IFG	fasting glucose	Glargine	Repaglinide	diabetic retinopathy
Type II diabetes	IGT	fasting plasma glucose	Glucophage	Rosiglitazone	diabetic skin spots
Type II DM	impaired fasting glucose	fingerstick glucose test	Glulisine	Ryzodeg	End stage renal disease
www diabetes org	impaired glucose tolerance	FPG	Glumetza	saxagliptin	ESRD
diabetes lose weight	impaired glycemic control	glucometers	glyset	SGLT2 inhibitors	foot ulcers
borderline diabetes	impaired insulin secretion	glucose meter	Humalog	short acting insulin	gangrene
diabetes kit	increased glucose production	glucose monitor	Humulin	Sitagliptin	gastroparesis
onset diabetes	insulin	glucose test strip	Humulin N	starlix	hypoglycemia
diabetes magazines	insulin deficiency	glucose tolerance test	inhaled insulin	SymlinPen	insulin reaction
herbal treatment for diabetes	insulin receptor	HbA1c	injection site rotation	thiazolidinediones	insulin shock
neuropathy	nocturnal diarrhea	peripheral neuropathy			

Appendix B

Cosine Similarity

Function cos(*q, d*) is the cosine similarity of *q* and *d*. Parameter *q* represents a search keyword or a comment. Parameter *d* represents a video title or description.

$$\cos(q,d) = \frac{\sum_{i=1}^{|V|} q_i d_i}{\sum_{i=1}^{|V|} q_i^2 \sqrt{\sum_{i=1}^{|V|} d_i^2}}$$

Where q_i is the term frequency of term i in the search keyword or comment, d_i is the term frequency of term i in the video title or description, V is the unique words in q and d, and $|V|$ is the total number of unique words in q and d.

Appendix C

Sample Videos with High or Low Medical Information

Video ID	Video Title	Rating	Explanation
Jk5IXhB9ZXM[1]	Diabetes Rx May Do Double Duty for Parkinson's Disease	High medical information	This video summarizes a research study from London School of Hygiene & Tropical Medicine (LSHTM) about unintended beneficial effect of glitazone, a diabetes medication on preventing Parkinson's disease (PD). This video includes medical information about a diabetes treatment and is presented by a professional reporter.
FnIf3OPDBFM[2]	Willie reduced insulin down from 72 units of insulin per day to 0 on Virta	Low medical information	In the video, patient Willie talked about a program, VirtaHealth, to reduce his insulin dosage. However, there are no specifics about what the program does, and how it reduces the insulin dosage for Willie.
QfRpnA3CKos[3]	Amylin Pharmaceuticals Increases Speed to Insight with 100 Million Rows	Low medical Information	Although this video shows up in the result when searching "diabetes treatment", it is about how to use Tableau to support data analysis in diabetes treatment development. Patient, medicines, and doctors are not discussed in the video.
QfE4CBG90tM[4]	Type 2 Diabetes and Prediabetes Fast Facts: Risk Factors	High medical information	This video discusses the risk factors of diabetes. It categorizes them into factors patients cannot change and factors patients can change. It also provides actionable advices on how to reduce the risk.
L2dpW_bFRc[5]	Metabolic Syndrome and Plant-Based Diets	High medical information	In the video, a doctor is promoting plant-based diet to diabetes patients. He references multiple research papers about metabolism in the video to support his viewpoint.
t0JF2Gbv1XQ[6]	Diabetes Nutrition – How Meghan Eats 300 Grams of Carbs with Excellent Blood Glucose Control T1D	Low medical information	This video is a conversation based interview from a diabetes patient Meghan with two persons running a plant-based diet program. Meghan causally shared what she eats on daily basis and two other persons are trying to promote their program. The amount of medical information is low because the only relevant part is Meghan's diet.

[1]https://www.youtube.com/watch?v=Jk5IXhB9ZXM
[2]https://www.youtube.com/watch?v=FnIf3OPDBFM
[3]https://www.youtube.com/watch?v=QfRpnA3CKos
[4]https://www.youtube.com/watch?v=QfE4CBG90tM
[5]https://www.youtube.com/watch?v=_L2dpW_bFRc
[6]https://www.youtube.com/watch?v=t0JF2Gbv1XQ